

Cross-domain Depth Estimation Network for 3D Vessel Reconstruction in OCT Angiography

No Author Given

No Institute Given

Abstract. Optical Coherence Tomography Angiography (OCTA) has been widely used by ophthalmologists for decision-making due to its superiority in providing capillary details. Many of the OCTA imaging devices used in clinic provide high-quality 2D *en face* representations, while their 3D data quality are largely limited by low signal-to-noise ratio and strong projection artifacts, which restrict the performance of depth-resolved 3D analysis. In this paper, we propose a novel 2D-to-3D vessel reconstruction framework based on the 2D *en face* OCTA images. This framework takes advantage of the detailed 2D OCTA depth map for prediction and thus does not rely on any 3D volumetric data. Based on the data with available vessel depth labels, we first introduce a network with structure constraint blocks to predict the depth of blood vessels in other cross-domain *en face* OCTA data with unavailable labels. Afterwards, a depth adversarial adaptation module is proposed for better unsupervised cross-domain training, since images captured using different devices may suffer from varying image contrast and noise levels. Finally, vessels are reconstructed in 3D space by utilizing the estimated depth map and 2D vascular information. Experimental results demonstrate the effectiveness of our method and its potential to guide subsequent vascular analysis in 3D domain.

Keywords: OCTA · domain adaptation · 3D vessel reconstruction.

1 Introduction

Optical Coherence Tomography Angiography (OCTA) is a novel 3D imaging technique that has the ability to acquire rich blood flow details at capillary-level in a short time, as shown in Fig. 1 (a). Compared with fluorescein angiography (FA) that is traditionally used to detect ocular diseases, OCTA is non-invasive and does not expose patients with side effects such as nausea or anaphylaxis. As a result, this budding technology has shown considerable potential in the diagnosis of various eye-related diseases such as diabetic retinopathy (DR), age related macular degeneration (AMD), and glaucoma [1,2].

Recently, several studies have been conducted to analyze retinal vessels in OCTA images. Hwang *et al.* [4] used a quantification method to analyze the correlation between OCTA features and DR. Arrigo *et al.* [5] applied OCTA images for the detection of the choroidal neovascularization area, in participants with

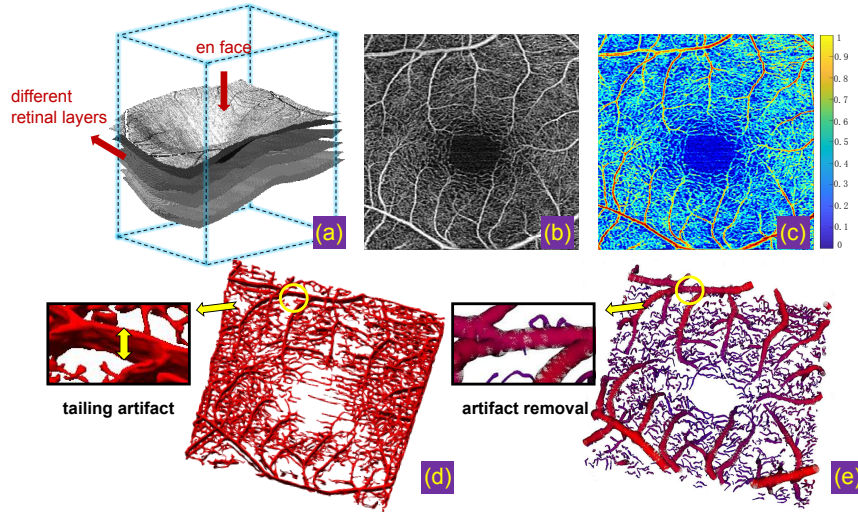


Fig. 1. Visualization of (a) an example 3D OCTA original volume. (b) 2D *en face* angiogram. (c) Color encoded depth map of (b). (d) Vessel 3D reconstruction by an existing method [3], and (e) reconstruction result by our method.

AMD. Xie *et al.* [6] estimated the vascular topology, and classified the retinal vessels into artery and vein from *en face* image. Some studies [7,8,9] have developed automatic methods for vessel segmentation in OCTA images. However, all the above-mentioned works are only performed on 2D *en face* angiograms (Fig. 1 (b)), and the 3D spatial information of the vessels are not utilized.

3D analysis and visualization of blood vessels can provide additional spatial information that is not available in 2D images, and are very useful for observing vascular changes [10,11,12]. Zhang *et al.* [3] built a 3D microvascular shape modeling framework based on the original volume of OCTA. Although this work provides a reliable 3D visualization and effective vascular analysis framework, it still suffers from projection artifacts as shown in Fig. 1 (d), and requires direct processing of 3D data. 3D vessel reconstruction from an OCTA volume faces several challenges: overlaps during the projection, complex topological structures, poor contrast, and high computational demand. Some public datasets such as ROSE [9] does not provide 3D data which makes it impossible to achieve 3D reconstruction of vessels through existed methods. Consequently, analysis and visualization of the OCTA images in 3D space still remain a challenge.

With the development of OCTA imaging technology, the dedicated depth-resolved information becomes available in current devices such as the CIRRUS HD-OCT 5000 System (Carl Zeiss Meditec Inc., USA), which is equipped with an *AngioPlexTM* OCT Angiography software with an additional color-coded *en face* OCTA image (we refer it as a depth map in this paper) as shown in Fig. 1 (c). The depth map is intrinsically a combination of the retinal layers shown in Fig. 1 (a). Similar to range images in computer vision, a depth image refers to an image whose pixel value is the distance from the imaging sensor to each position in the scene. For a depth map obtained by CIRRUS HD-OCT 5000, the

red color indicates that the vessels are closer to the imaging sensor whilst blue represents vessels that are further away.

In this work, we propose a framework to reconstruct the 3D structure of blood vessels from 2D OCTA images, while 3D data is utilized for verification. Firstly, we introduce a cross-domain depth estimation network with structure constraint blocks (SCB) in a depth adversarial adaptation (DAA) module, so as to reduce domain discrepancies and obtain improved results. Secondly, the 3D vessel are reconstructed by integrating 2D vascular information with the predicted depth maps. The comparison and ablation study based on two datasets demonstrate the effectiveness of our method. Compared to our previous work [], this work employed the concept of domain adaptation.... is the first attempt to achieve 3D microvascular visualization from multi-domain 2D OCTA data including images acquired by Cirrus HD-OCT and RTVue XR Avanti SD-OCT.

This framework can be generalized to 2D *en face* OCTA images from a variety of imaging devices.

To our best knowledge, this is the first work to generate and visualise 3D vessel structures from 2D *en face* OCTA angiograms based on the depth map prediction. It is worth noting that while achieving reconstruction accuracy equal to that of the existing state-of-the-art method, our method can also avoid the interference of projection artifact as illustrated in Fig. 1 (e), and is not limited by data domains.

2 Proposed Method

In this section, we detail the proposed 3D vessel reconstruction method in two steps: depth map prediction and 3D reconstruction.

2.1 Cross-domain Depth Estimation Network

Network Architecture: In view of the excellent performance of the U-shaped network [13] in biomedical images, we apply it as the backbone of our model. However, direct skipping connections between encoder and decoder in [13] will transmit redundant information. In our depth prediction task, the accuracy of depth for image regions with blood vessels is particularly important for the subsequent reconstruction, subject to very complex vascular topology structural constraints. To this end, we employ a structure branch to process only vessel-related information by means of our carefully designed SCB and local supervision.

As illustrated in Fig. 2, SCB is enforced after each encoder block and connected to the corresponding decoder block. Let e_t ($t \in 2, \dots, 5$) denotes the output of t_{th} encoder block, and $s_{\tilde{t}}$ denotes the corresponding intermediate representations of the structure branch. We first obtain an attention map a_{t-1} by concatenating e_t and $s_{\tilde{t}}$ followed by convolution, batch normalization and non-linear activation layers. Given the attention map a_{t-1} , an element-wise product is applied between e_t and a_{t-1} to acquire a weighted map. Note that upsampling is employed on e_t before concatenation to ensure that e_t and $s_{\tilde{t}}$ have the same size. Since they contain rich edge information, the low-level features from the first block of the encoder are used to obtain the initial weighted map. Intuitively,

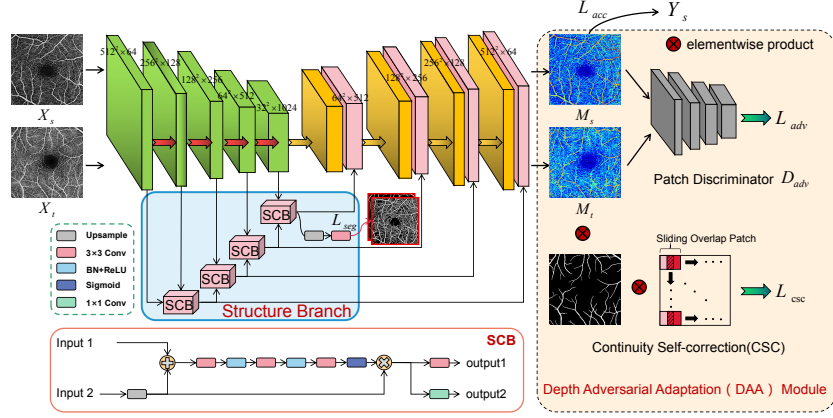


Fig. 2. Architecture of the proposed cross-domain depth estimation network.

a_{t-1} can be seen as an attention map that weights more important areas with vasculature. The filtered feature maps by SCB (i.e., output2 in Fig. 2) are cascaded with the corresponding decoder feature maps to provide refined structure-related information, and the output of the last SCB is subject to upsampling and convolution operations to finally obtain the vessel prediction map.

To optimize the capability for domain adaptation, a DAA module is then attached to narrow the domain discrepancies. Formally, the source domain and target domain are denoted as S and T, respectively. The input images from S are denoted as $\mathcal{X}_s \in \mathbb{R}^{C \times H \times W}$ while those from T are denoted as $\mathcal{X}_t \in \mathbb{R}^{C \times H \times W}$, where C is the channel dimension, H and W are spatial dimensions. \mathcal{M}_s and \mathcal{M}_t represent the predictions of S and T, while \mathcal{Y}_s denotes the label of the source image. We first feed \mathcal{M}_s and \mathcal{M}_t into the patch discriminator D_{adv} for adversarial training, so as to make the distribution of \mathcal{M}_t closer to that of \mathcal{M}_s . In addition, we make full use of the prior knowledge of vascular continuity and employ continuity self-correction (CSC) to obtain final outputs.

CSC first multiplies \mathcal{M}_t , and its segmented image includes a pixel by pixel depth map \mathcal{M}_t in effective vessel areas. As an ordered topology, in fact the depths of adjacent blood vessels are continuous in physical space. Based on this important prior knowledge, we then use a sliding overlapping patch to traverse \mathcal{M}_t . In each infinitesimal patch size, we minimize the variance of the depth of blood vessels so that the continuity of vascular depth can be guaranteed, and the outliers can be corrected. The overlapping part is set to make adjacent blood vessels in the same vascular tree continuous in depth.

Loss Functions: In this work, the loss function consists of four parts \mathcal{L}_{seg} , \mathcal{L}_{acc} , \mathcal{L}_{adv} and $\mathcal{L}_{continuity}$, to be defined below. We first use the cross-entropy (CE) loss \mathcal{L}_{seg} on the predicted segmentation maps $pred_{seg}$:

$$\mathcal{L}_{seg} = \mathcal{L}_{CE}(pred_{seg_s}, gt_{seg_s}) + \mathcal{L}_{CE}(pred_{seg_t}, gt_{seg_t}), \quad (1)$$

where $gt_{seg_s} \in \mathbb{R}^{C \times H \times W}$ and $gt_{seg_t} \in \mathbb{R}^{C \times H \times W}$ denote the groundtruths of the source domain and target domain, respectively. In order to improve the

accuracy of the predicted depth map at both pixel- and overall structure- level, we combine MSE and SSIM [14] loss as \mathcal{L}_{acc} between \mathcal{M}_s and \mathcal{Y}_s :

$$\mathcal{L}_{acc} = \mathcal{L}_{MSE}(\mathcal{M}_s, \mathcal{Y}_s) + \mathcal{L}_{SSIM}(\mathcal{M}_s, \mathcal{Y}_s). \quad (2)$$

For the adversarial training, we use the adversarial loss \mathcal{L}_{adv} as:

$$\mathcal{L}_{adv} = \min_G \max_D \mathbb{E}[\log D(\mathcal{M}_s)] + \mathbb{E}[\log(1 - D(G(\mathcal{X}_t)))], \quad (3)$$

where G is the generator (i.e. $G(\mathcal{X}_t) = \mathcal{M}_t$) and D is the patch discriminator D_{adv} . \mathcal{L}_{adv} is designed to train the network and to deceive D_{adv} by maximizing the probability of T being considered as S .

Additionally, a continuity self-correction loss \mathcal{L}_{csc} is defined to ensure the continuity of the depth of adjacent blood vessels:

$$\mathcal{L}_{csc} = \text{Var}[\sum_i \text{Patch}(i)], \quad (4)$$

in which $\text{Var}[\cdot]$ denotes the variance of input \cdot and $\text{Patch}(i)$ denotes the i_{th} sliding patch of \mathcal{M}_t (the patch size is set to 8×8 and the sliding step is 4 in this paper). The final loss function \mathcal{L}_{total} of the proposed method is thus:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{acc} + \lambda_2 \mathcal{L}_{seg} + \lambda_3 \mathcal{L}_{adv} + \lambda_4 \mathcal{L}_{csc}, \quad (5)$$

where λ_1 , λ_2 , λ_3 and λ_4 are empirically set in this paper to 100, 7, 1 and 1, respectively.

2.2 3D vessel reconstruction via depth map

The 3D reconstruction of blood vessels can be regarded as a mapping problem from 2D to 3D with the availability of 2D segmented vessels and depth maps. In this step, we remove artifacts in OCTA, so that the final reconstructed surface can be used for subsequent 3D feature extraction and analysis.

3D Centerline Point Cloud Extraction Given an OCTA segmentation image, the vessel centerlines as well as diameters can be extracted using the skeletonization method [15], and the bifurcation points of the vessel network can be detected by locating intersection points (pixels with more than two neighbours). All the intersection points and their neighbours may then be removed from the centreline map, in order to obtain an image with clearly separated vessel segments. The depth of each centerline point depends on the corresponding position of its depth map and is shifted downward by a distance of radius of vessel. Therefore, a 3D point cloud comprising the centerline points is obtained, as shown in Fig. 3 (a), where adjacent segments are linked according to the topology consistency using bilinear interpolation.

Vessel Surface Reconstruction Using the previously obtained centerline point cloud, a center-to-surface method is employed to convert centerline to surface. Taking each center point as the centroid of a circle and the line between two adjacent points as the normal vector, a circle is then generated with a predefined radius .

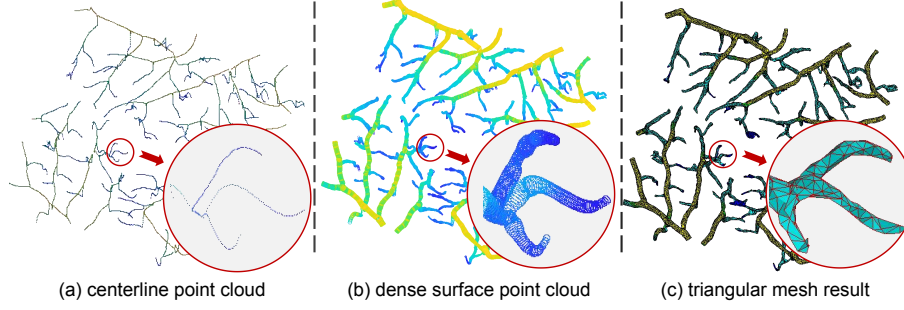


Fig. 3. Illustration of vascular reconstruction from point cloud to mesh surface.

Next, we obtain sampling points at equal intervals for each circle, and traverse all segments to obtain a dense surface point cloud, as shown in Fig. 3 (b). Furthermore, we adapted a constrained Poisson-disk sampling [16] to refine the surface so as to obtain a more uniform point cloud, while ensuring that the objects of a certain size can be distributed according to the sampling scheme, without overlapping. Finally, a Ball-pivoting algorithm [17] is utilized to generate the final 3D triangular mesh, as shown in Fig. 3 (c).

3 Experiments

In this section, we validate the performance of the proposed method in terms of depth map prediction and 3D vessel reconstruction.

3.1 Datasets and Evaluation Metrics

Datasets: Two OCTA datasets each containing 80 pairs of images were obtained from a CIRRUS HD-OCT 5000 system (Carl Zeiss Meditec Inc., USA) and an RTVue XR Avanti system (Optovue, Fremont Inc., USA). The former was selected as the source domain as it is associated with depth maps, the latter was selected as the target domain. The scan area was $3mm \times 3mm \times 2mm$ centered on the fovea, and both of the source and target images used in this work were resized to 512×512 pixels.

56 of 80 images in the source domain were used for training, and the rest for testing. 56 of 80 images in the target domain were used for training, and the rest for validation. It is worth noting that a state-of-the-art OCTA vessel segmentation model, OCTA-Net [9], was used to extract vessels from the training set, and an image analysis expert further refined the vessel segmentation results as the ground truth, i.e., gt_{seg_s} in Eq.(1). The target domain data include the corresponding segmentation annotations, i.e., gt_{seg_t} in Eq.(1).

The method utilized in [3] was employed to obtain a 3D vessel segmentation result from the raw 3D volume. A centerline point cloud was obtained by taking the upper surface of the vessel as the depth, and moving downward with a distance of the corresponding radius. Finally, the aforementioned method was applied to obtain the surface as the ground truth. For comparison purposes, the

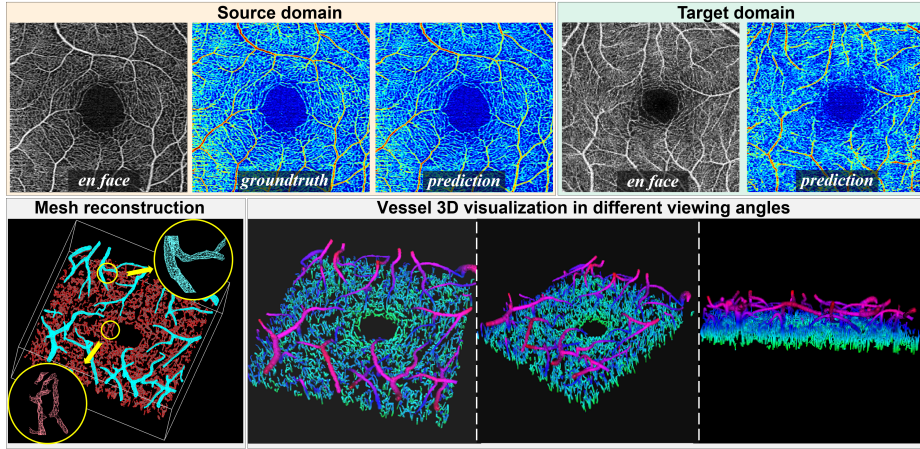


Fig. 4. Illustrative results of depth prediction in terms of source and target domain images, and 3D vessel reconstruction.

ground truth point cloud was mapped to the same spatial depth range (0-255) as that reconstructed via the predicted depth map in the target domain.

Metrics: A total of five metrics were employed to verify the performance of the proposed method in both the source and target domains. For the source domain, the accuracy (ACC) metric δ [18] was used to validate the predicted depth map with its ground truth: $\delta = \max(\frac{D_i}{D_i^*}, \frac{D_i^*}{D_i}) < T$, where D_i and D_i^* are the estimated depth and the corresponding depth of the i -th pixel of the ground truth, respectively. As suggested in [19], a threshold $T = 1.25$ was used in this metric. The Absolute Relative Difference (ARD) and Root Mean Squared Error (RMSE), two most commonly-used metrics in evaluating monocular image depth estimation, were also used in this work. For the target domain, Chamfer Distance (CD)[20] and Hausdorff Distance(HD)[21] were utilized to measure the similarity between two sets of points.

3.2 Experimental Results

Fig. 1(d-e) illustrate the 3D vessel reconstruction results by method [3] and the proposed method, where we can observe the projection artifact along big vessels. This problem has been certainly addressed based our method, and thus it will be more useful to ophthalmologists in the diagnosis and treatment of retinal diseases. More visual results can be found in bottom row of Fig. 4. To further verify the superiority of our method, the following comparative and ablation studies were also carried out.

Comparison with previous methods Table 1 reports the performances of the proposed method compared with state-of-the-art depth predicting methods [19,22,23,24] and the most commonly-used domain adaptation algorithms [25,26,27]. For the source domain, it can be seen that our method is able to

Table 1. Comparison of the proposed method with the state-of-the-art methods.

Methods	Source Domain			Target Domain	
	ACC	ARD	RMSE	CD	HD
Eigen <i>et al.</i> [22]	0.596	0.310	0.343	5.266	6.913
Laina <i>et al.</i> [19]	0.925	0.138	0.213	3.452	6.874
Chen <i>et al.</i> [23]	0.973	0.114	0.138	2.921	5.147
Yu <i>et al.</i> [24]	0.971	0.058	0.107	2.624	3.502
Tzeng <i>et al.</i> [25]	0.795	0.241	0.293	1.892	4.730
Yue <i>et al.</i> [26]	0.983	0.056	0.098	1.549	4.088
Li <i>et al.</i> [27]	0.971	0.125	0.187	1.463	4.279
Proposed method	0.984	0.049	0.096	1.375	4.107

Table 2. Ablation study of the proposed method.

SCB	D_{adv}	CSC	CD	HD
			3.413	6.610
	✓		1.897	4.662
✓	✓		1.764	4.575
✓	✓	✓	1.375	4.107

generate a depth map similar to the ground truth as shown in top row of Fig. 4: this is confirmed by ACC, ARD and RMSE in Table 1. For the target domain, both CD and HD show the similarity between the reconstructed point cloud and corresponding ground truth. Overall, the proposed method outperforms the previous methods in terms of almost all the metrics by significant margins for both depth prediction and domain adaptation.

Ablation study To understand the roles of different components in the proposed method, an ablation study is performed. Table 2 summarizes the results of the proposed method with different combinations of components. D_{adv} significantly decreases the value of CD from 3.413 to 1.897, and HD from 6.610 to 4.662. The addition of SCB and CSC, separately and in combination, also effectively improves the network capacity. A combination of the SCB, D_d and CSC provides significant improvements, which confirms that the proposed method produces the best results.

4 Conclusion

In this work, we have proposed a novel method for the 3D reconstruction of vessels in OCTA images via depth map estimation, which is suitable for OCTA images obtained from a variety of devices. The remarkable significance of this work is that it successfully demonstrates the effective use of 2D OCTA *en face* angiograms alone for 3D vessel reconstruction, and that it is applicable to images in two different data domains. Moreover, it effectively solves the projection artifact problem. The high evaluation performance demonstrates the effectiveness of our method both qualitatively and quantitatively. It reveals considerable potential to exploring the subsequent vessel analysis in 3D space, and to assist clinical research in the future.

5 Acknowledgment

This work was supported by

References

1. De Carlo, T.E., Romano, A., Waheed, N.K., Duker, J.S.: A review of optical coherence tomography angiography (octa). *International Journal of Retina and Vitreous* **1**(1) (2015) 1–15
2. Kashani, A.H., et al.: Optical coherence tomography angiography: A comprehensive review of current methods and clinical applications. *Progress in Retinal and Eye Research* **60** (2017) 66–100
3. Zhang, J., et al.: 3d shape modeling and analysis of retinal microvasculature in oct-angiography images. *IEEE Transactions on Medical Imaging* **39**(5) (2019) 1335–1346
4. Hwang, T.S., et al.: Automated quantification of capillary nonperfusion using optical coherence tomography angiography in diabetic retinopathy. *JAMA Ophthalmology* **134**(4) (2016) 367–373
5. Arrigo, A., et al.: Advanced optical coherence tomography angiography analysis of age-related macular degeneration complicated by onset of unilateral choroidal neovascularization. *American Journal of Ophthalmology* **195** (2018) 233–242
6. Xie, J., et al.: Classification of retinal vessels into artery-vein in oct angiography guided by fundus images. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer (2020) 117–127
7. Eladawi, N., et al.: Automatic blood vessels segmentation based on different retinal maps from octa scans. *Computers in Biology and Medicine* **89** (2017) 150–161
8. Mou, L., et al.: Cs-net: channel and spatial attention network for curvilinear structure segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer (2019) 721–730
9. Ma, Y., et al.: Rose: A retinal oct-angiography vessel segmentation dataset and new model. *IEEE Transactions on Medical Imaging* **40**(3) (2021) 928–939
10. Nesper, P.L., Soetikno, B.T., Treister, A.D., Fawzi, A.A.: Volume-rendered projection-resolved oct angiography: 3d lesion complexity is associated with therapy response in wet age-related macular degeneration. *Investigative Ophthalmology & Visual Science* **59**(5) (2018) 1944–1952
11. Zhao, Y., Zheng, Y., Liu, Y., Zhao, Y., Luo, L., Yang, S., Na, T., Wang, Y., Liu, J.: Automatic 2-d/3-d vessel enhancement in multiple modality images using a weighted symmetry filter. *IEEE Transactions on Medical Imaging* **37**(2) (2017) 438–450
12. Sarabi, M.S., et al.: 3d retinal vessel density mapping with oct-angiography. *IEEE Journal of Biomedical and Health Informatics* **24**(12) (2020) 3466–3479
13. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer (2015) 234–241
14. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**(4) (2004) 600–612
15. Bankhead, P., Scholfield, C.N., McGeown, J.G., Curtis, T.M.: Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PloS One* **7**(3) (2012) e32435

16. Corsini, M., Cignoni, P., Scopigno, R.: Efficient and flexible sampling with blue noise properties of triangular meshes. *IEEE Transactions on Visualization and Computer Graphics* **18**(6) (2012) 914–924
17. Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., Taubin, G.: The ball-pivoting algorithm for surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics* **5**(4) (1999) 349–359
18. Ladicky, L., Shi, J., Pollefeys, M.: Pulling things out of perspective. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2014) 89–96
19. Laina, I., Rupprecht, C., Belagiannis, V., Tombari, F., Navab, N.: Deeper depth prediction with fully convolutional residual networks. In: *International Conference on 3D Vision (3DV)*, IEEE (2016) 239–248
20. Borgefors, G.: Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing* **34**(3) (1986) 344–371
21. Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J.: Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**(9) (1993) 850–863
22. Eigen, D., Fergus, R.: Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. In: *IEEE International Conference on Computer Vision*. (2015) 2650–2658
23. Chen, W., Fu, Z., Yang, D., Deng, J.: Single-image depth perception in the wild. In: *International Conference and Workshop on Neural Information Processing Systems* (2016)
24. Yu, S., et al.: 3D Vessel Reconstruction in OCT-Angiography via Depth Map Estimation. *arXiv e-prints* (2021) arXiv:2102.13588
25. Tzeng, E., Hoffman, J., Saenko, K., Darrell, T.: Adversarial discriminative domain adaptation. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2017) 7167–7176
26. Yue, X., Zhang, Y., Zhao, S., Sangiovanni-Vincentelli, A., Keutzer, K., Gong, B.: Domain randomization and pyramid consistency: Simulation-to-real generalization without accessing target domain data. In: *IEEE/CVF International Conference on Computer Vision*. (2019) 2100–2110
27. Li, S., et al.: Domain conditioned adaptation network. In: *AAAI Conference on Artificial Intelligence*. Volume 34. (2020) 11386–11393